Theoretical Computer Science 289 (2002) 997–1008

Theoretical Computer Science

www.elsevier.com/locate/tcs

Parameterized complexity of finding subgraphs with hereditary properties ☆

Subhash Khot[a], Venkatesh Raman[b,*]

[a] Department of Computer Science, Princeton University, NJ, USA

[b] The Institute of Mathematical Sciences, C.I.T. Campus, Chennai-600113, India

Accepted November 2001

Abstract

We consider the parameterized complexity of the following problem under the framework introduced by Downey and Fellows: Given a graph G, an integer parameter k and a nontrivial hereditary property Π, are there k vertices of G that induce a subgraph with property Π? This problem has been proved NP-hard by Lewis and Yannakakis. We show that if Π includes all trivial graphs but not all complete graphs or vice versa, then the problem is complete for the parameterized class $W[1]$ and is fixed parameter tractable otherwise. Our proofs of both the tractability and hardness involve nontrivial use of the theory of Ramsey numbers.

Keywords: Parameterized complexity; Hereditary properties; Ramsey numbers

1. Introduction

Many computational problems involve two parts associated with the problem input. For example, we may be concerned with finding a vertex cover or a clique of size k in a graph G on n vertices. While both these problems are (equally) classified as NP-complete (when k is part of the input) by traditional complexity theory, the parameter k contributes to the complexity of the problems in qualitatively different ways. The parameterized versions of the Vertex Cover (and also for example, Undirected Feedback Vertex Set) problem can be solved in $O(f(k)n^{\alpha})$ time where α is a constant

☆ Part of this work was done while the first author was at IIT Bombay and visited the Institute of Mathematical Sciences, Chennai. Part of the second author's research was supported by a DST-DAAD grant.

* Corresponding author.

E-mail addresses: khot@cs.princeton.edu (S. Khot), vraman@imsc.ernet.in (V. Raman).

PII: S0304-3975(01)00414-5

independent of k and f is an arbitrary function of k (against a naive $\Theta(n^{ck})$ algorithm for some constant c). This "good behavior", which is extremely useful in practice for small values of k, is termed *fixed parameter tractability* in the theory introduced by Downey and Fellows [3–5].

On the other hand, for problems like CLIQUE (and also for example, DOMINATING SET), the best known algorithms for the parameterized version has complexity $\Theta(n^{ck})$ for some constant c. CLIQUE problem is known to be hard for the parameterized complexity class $W[1]$ (and Dominating Set, hard for $W[2]$) and is considered unlikely to be *fixed parameter tractable* (denoted by FPT). See [5] for the definitions and more on parameterized complexity theory. In this paper, we investigate the parameterized complexity of finding induced subgraphs of an arbitrary, but fixed, nontrivial hereditary property in a given graph.

A graph property Π is a collection of graphs. A graph property Π is nontrivial if it holds for at least one graph and does not include all graphs. A nontrivial graph property Π is said to be *hereditary* if $G \in \Pi$ implies that every *induced subgraph* of G is also in Π. A graph property is said to be *interesting* [11] if the property is true (as well as false) for infinite families of graphs. Lewis and Yannakakis [11] (see also [8]) showed that if Π is a nontrivial and interesting hereditary property, then it is NP-hard to decide whether for a given graph G, k vertices can be deleted to obtain a graph $G' \in \Pi$.

Cai [1] considered the parameterized version of this problem. He has shown that a more general graph modification problem $\Pi(i, j, k)$ defined below is fixed parameter tractable (FPT) for a nontrivial hereditary property Π with a finite forbidden set. We will now define this notion of a forbidden set for a hereditary property. For a hereditary property Π, let $\mathscr{F}$ be the family of graphs not having the property. The set of minimal members (minimal with respect to the operation of taking induced subgraphs) of $\mathscr{F}$ is called the *forbidden set* for the property Π. For example, the collection of all bipartite graphs is a hereditary property whose forbidden set consists of all odd cycles. Conversely, given any family $\mathscr{F}$ of graphs, we can define a hereditary property by declaring its forbidden set to be the set of all minimal members of $\mathscr{F}$.

The problem considered by Cai is given below. Let Π be a nontrivial hereditary property with a finite forbidden set.

Given: A simple undirected graph $G = (V, E)$ with vertex set V and edge set E where $|V| = n$ and $|E| = m$.

Parameter(s): Integers i, j, k.

Question. Are there sets $V' \subseteq V$, $E' \subseteq E$ and $E'' \subseteq E^c$ with $|V'| = i$, $|E'| = j$ and $|E''| = k$ such that $G - V' - E' \cup E''$ is in Π? (Here E^c is the set of edges in the complement of the graph.)

The problem $\Pi(k, 0, 0)$ is the node deletion problem addressed by Lewis and Yannakakis. The parameterized complexity of the problem when Π is a hereditary property with an infinite forbidden set is still open. In this paper, we address the parametric dual of the node-deletion problem defined below. Given any property Π, the problem $P(G, k, \Pi)$ is defined as follows.

Given: A simple undirected graph $G = (V, E)$.

Parameter: An integer $k \leqslant |V|$.

Question. Is there a subset $V' \subseteq V$ with $|V'| = k$ such that the subgraph of G induced by V', $G[V']$ is in Π?

This problem is the same as $\Pi(|V| - k, 0, 0)$ problem (i.e. can we delete all but k vertices of G to get a graph in property Π) and hence NP-hard. However, the parameterized complexity of this problem does not follow from the complexity of the problem addressed by Cai even for properties having a finite forbidden set. This is because the NP-hard reduction reduces a general instance (G, i, j, k) of the $\Pi(i, j, k)$ problem to the instance $P(G, n-k, \Pi)$ where n is $|V(G)|$. This is not a parameterized reduction as the parameter of the reduced instance can be a function only of the original parameter k (and not of n) in a parameterized reduction.

We prove that if Π includes all trivial graphs but not all complete graphs, or vice versa, then the problem $P(G, k, \Pi)$ is $W[1]$-complete. (By a trivial graph, we mean a graph with no edges.) The proof is by a parametric reduction from the Independent Set problem. To show that the problem is in $W[1]$, we reduce it to the k-step Turing machine halting problem [5]. If Π includes all trivial graphs and all complete graphs, or excludes some trivial graph and some complete graph, then we show that the problem is fixed parameter tractable.

Cai's result coupled with ours strengthens the observation that the parametric dual problems usually have complimentary parameterized complexity. This phenomenon has been observed in many other parameterized problems as well. For a graph $G(V, E)$, finding a vertex cover of size k is FPT whereas finding an independent set of size k (or a vertex cover of size $|V| - k$) is $W[1]$-complete. Given a boolean 3-CNF formula with m clauses, finding an assignment to its variables that satisfies at least k clauses is FPT, whereas finding an assignment that satisfies at least $(m-k)$ clauses (i.e. all but at most k clauses) is known to be $W[P]$-hard [3,12] (k is the parameter in both these problems). The k-Irredundant Set problem is $W[1]$-hard whereas Co-Irredundant set or $(n-k)$ Irredundant Set problem is FPT [6]. Our result adds another problem to this list. There are also examples of parameterized dual problems that do not have such complimentary parameterized complexity [7].

The next section gives some definitions mostly pertaining to the notion of parameterized complexity. Section 3 deals with the hereditary properties for which the problem is fixed parameter tractable, and Section 4 proves the $W[1]$-hardness result for the remaining hereditary properties. Section 5 concludes with some remarks and open problems.

2. Definitions

Throughout the paper, by a graph we mean an undirected graph with no loops or multiple edges. By a nontrivial graph, we mean a graph with at least one edge. Given a graph G and $A \subseteq V(G)$, by $G[A]$ we mean the subgraph of G induced by vertices in A. For two graphs H and G, we use the notation $H \subseteq G$ to mean that H is isomorphic to an induced subgraph of G. For the graph properties Π we will be concerned with in this paper, we assume that Π is decidable; i.e. given a graph G on n vertices, one can decide whether or not G has property Π in $f(n)$ time for some function f of n.

We have already introduced the notion of fixed parameter tractability. We give below some notion of the complexity classes associated with the hard problems. See [5] for more details.

A parameterized language L is a subset of $\Sigma^* \times \mathbb{N}$ where Σ is some finite alphabet and $\mathbb{N}$ is the set of all natural numbers. For $(x,k) \in L$, k is the parameter. We say that a parameterized problem A reduces to a parameterized problem B, if there is an algorithm Φ which transforms (x,k) into $(x', g(k))$ in time $f(k)|x|^{\alpha}$ where $f,g: \mathbb{N} \to \mathbb{N}$ are arbitrary functions and α is a constant independent of k, so that $(x,k) \in A$ if and only if $(x', g(k)) \in B$. The essential property of parametric reductions is that if A reduces to B and if B is FPT, then so is A.

Let F be a family of boolean circuits with *and*, *or* and *not* gates. We allow that F may have many different circuits with a given number of inputs. Let the weight of a boolean vector be the number of ones in the vector. To F we associate the parameterized circuit problem $L_F = \{(C,k)$: C accepts an input vector of weight $k\}$. Let the *weft* of a circuit be the maximum number of gates with fan-in more than two, on an input–output path in the circuit.

A parameterized problem L belongs to $W[t]$ if L reduces to the parameterized circuit problem $L_{F(t,h)}$ for the family $F(t,h)$ of boolean circuits with the weft of the circuits in the family bounded by t, and the depth of the circuits in the family bounded by a constant h. This naturally leads to a completeness program based on a hierarchy of parameterized problem classes:

$$FPT \subseteq W[1] \subseteq W[2] \subseteq \cdots .$$

In the same way as an NP-complete decision problem is considered unlikely to have a polynomial time solution, $W[1]$-hardness is considered an evidence of the fact that a parameterized problem is unlikely to be in FPT. Parameterized independent set, clique and k-step halting Turing machine problem are some of the canonical $W[1]$-complete problems.

3. Hereditary properties that are FPT

In this section, we identify hereditary properties Π for which the $P(G,k,\Pi)$ problem is fixed parameter tractable.

Lemma 1. *If a hereditary property Π includes all trivial graphs and all complete graphs, or excludes some trivial graphs as well as some complete graphs, then the problem $P(G,k,\Pi)$ is fixed parameter tractable.*

Proof. For any positive integers p and q, there exists a minimum number $R(p,q)$ (the Ramsey number) such that any graph on at least $R(p,q)$ vertices contains either a clique of size p or an independent set of size q. It is well-known that $R(p,q) \leqslant \binom{p+q-2}{q-1}$ [10].

Assume that Π includes all complete graphs and trivial graphs. For any graph G with $|V(G)| \geqslant R(k,k)$, G contains either a clique of size k or an independent set of size k. Since all independent sets and all cliques have property Π, the answer to the problem $P(G,k,\Pi)$ in this case is "yes".

When $|V(G)|\leqslant R(k,k)$, we can use brute force by picking all k-elements subsets of $V(G)$ and checking whether the induced subgraph on the subset has property Π. This will take time at most $\binom{R(k,k)}{k}f(k)$ where $f(k)$ is the time to decide whether a given graph on k vertices has property Π. Thus the problem $P(G,k,\Pi)$ is fixed parameter tractable.

If Π excludes some complete graphs and some trivial graphs, let s and t, respectively, be the sizes of the smallest complete graph and the smallest trivial graph which do not have property Π. Since any graph with at least $R(s,t)$ vertices has either a clique of size s or an independent set of size t, no graph with at least $R(s,t)$ vertices can have property Π (since Π is hereditary). Hence any graph in Π has at most $R(s,t)$ vertices and hence Π contains only finitely many graphs. So if $k>R(s,t)$, then the answer to the $P(G,k,\Pi)$ problem is NO for any graph G. If $k\leqslant R(s,t)$, then check, for each k subset of the given vertex set, whether the induced subgraph on the subset has property Π. This will take time $\binom{n}{k}f(k)\leqslant Cn^{R(s,t)}$ for an n vertex graph, where C is the time taken to check whether a graph of size at most $R(s,t)$ has property Π. Since s and t depend only on the property Π, and not on k or n, and $k\leqslant R(s,t)$, the problem $P(G,k,\Pi)$ is fixed parameter tractable in this case also. □

We list below a number of hereditary properties Π (dealt with in [13]) each of which includes all trivial graphs and complete graphs, and hence for which the problem $P(G,k,\Pi)$ is fixed parameter tractable.

Corollary 2. *Given any simple undirected graph G, and an integer k, it is fixed parameter tractable to decide whether there is a set of k vertices in G that induces* (a) *a perfect graph*, (b) *an interval graph* (c) *a chordal graph*, (d) *a split graph*, (e) *an asteroidal triple free* (*AT-free*) *graph*, (f) *a comparability graph*, *or* (g) *a permutation graph*. (*See* [13] *or* [9] *for the definitions of these graphs.*)

4. Hereditary properties that are W[1]-complete

In this section, we show that the problem $P(G,k,\Pi)$ is $W[1]$-complete if Π includes all trivial graphs but not all complete graphs or vice versa.

For a graph G, let $\bar{G}$ denote the edge complement of G. For a property Π, let $\bar{\Pi}=\{\bar{G}\mid G \textit{ has property } \Pi\}$. We note that Π is hereditary if and only if $\bar{\Pi}$ is hereditary, and Π includes all trivial graphs but not all complete graphs if and only if $\bar{\Pi}$ includes all complete graphs but not all trivial graphs. Thus it suffices to prove $W[1]$-hardness when Π includes all trivial graphs, but not all complete graphs.

First we will show that the problem is in $W[1]$.

Lemma 3. *Let Π be a nontrivial decidable hereditary property. Then the problem $P(G,k,\Pi)$ is in $W[1]$.*

Proof. We will reduce the problem to the $W[1]$-complete problem *short Turing machine acceptance* [5], taking inspiration from the $W[1]$ membership proof of the Perfect Code problem [2]. The problem is defined below.

Input: A single tape nondeterministic Turing machine M having up to n transitions possible at each configuration, and a string x.

Parameter: A positive integer k.

Question. Does M have a computation path accepting x in at most k steps?

Let N be a deterministic Turing machine which, given a graph G on n vertices, and a list of k vertex names on its tape, decides whether or not the subgraph induced on the k vertices has property Π, in $f(k)$ time for some function f. Since we have assumed that the properties we deal with are recursive, such a deterministic Turing machine having, say, $g(k)$ transitions exists.

Now given a graph G on n vertices, the nondeterministic Turing machine we construct, will start from an empty tape and the initial state, and will nondeterministically select k vertices from the graph and write their names in a sequence on the tape. Then it will pass control to the deterministic Turing machine N which will verify that the k vertices selected induce a subgraph having property Π (N will also do some preprocessing to determine that the k vertices selected are distinct and will reject the input otherwise). So the composite nondeterministic Turing machine will have at most $nk + k^2 + g(k)$ transitions.

It is easy to see that the given graph G is an YES instance for the problem $P(G,k,\Pi)$ if and only if the nondeterministic Turing machine constructed will halt in $k+k^2+f(k)$ steps on the empty string. □

We now proceed to show that the problem is $W[1]$-hard.

4.1. Properties with finite forbidden sets

In this section, we will assume that the forbidden set of Π is finite. We first prove that if one of the graphs in the forbidden set of Π is a complete bipartite graph, then the problem $P(G,k,\Pi)$ is $W[1]$-hard.

Lemma 4. *Let Π be a hereditary property that includes all trivial graphs but not all complete graphs, and that has a finite forbidden set $\mathscr{F} = \{H_1, H_2, \ldots, H_s\}$. Assume that some H_i, say H_1 is a complete bipartite graph. Then the problem $P(G,k,\Pi)$ is $W[1]$-complete.*

Proof. In Lemma 3 we have shown that the problem is in $W[1]$.

Let Π be as specified in the lemma. Let $t = \max(|V_1|, |V_2|)$ where $V_1 \cup V_2$ is the bipartition of H_1. If $t = 1$, $H_1 = K_2$, and the given problem P is identical to the k-independent set problem, hence $W[1]$-hard. So assume $t \geqslant 2$. Note that $H_1 \subseteq K_{t,t}$. Let H_s be the clique of smallest size that is not in Π, hence in the forbidden set $\mathscr{F}$.

Now we will show that the problem is $W[1]$-hard by a parameterized reduction from the k-independent set problem which has been proved $W[1]$-complete [5]. Let G_1 be a graph in which we are interested in finding an independent set of size k_1. For every vertex $u \in G_1$ we take r independent vertices (r to be specified later) $u^1, \ldots, u^r$ in G. If (u, v) is an edge in G_1, we add all r^2 edges (u^i, v^j) in G. G has no other edges.

We claim that G_1 has an independent set of size k_1 if and only if G has rk_1 vertices that induce a subgraph with property Π (for an appropriate value of r).

Suppose G_1 has an independent set $\{u_i \mid 1 \leqslant i \leqslant k_1\}$ of size k_1. Then the set of rk_1 vertices $\{u_i^j \mid 1 \leqslant i \leqslant k_1, 1 \leqslant j \leqslant r\}$ is an independent set in G and hence has property Π.

Conversely, let S be a set of rk_1 vertices in G which induces a subgraph with property Π. This means that $G[S]$ does not contain any H_i, in particular it does not contain H_1. Group the rk_1 vertices according to whether they correspond to the same vertex in G_1 or not. Let $X_1, \ldots, X_h$, $Y_1, \ldots, Y_p$ be the groups and $u_1, \ldots, u_h$, $v_1, \ldots, v_p$ be the corresponding vertices in G_1 such that $|X_i| \geqslant t$ $\forall i$ and $|Y_j| < t$ $\forall j$. Observe that $\{u_1, \ldots, u_h\}$ must be independent in G_1 because if we have an edge (u_i, u_j), $H_1 \subseteq K_{t,t} \subseteq G[X_i \cup X_j] \subseteq G[S]$, a contradiction. If $h \geqslant k_1$ we have found an independent set of size at least k_1 in G_1. Therefore, assume that $h \leqslant k_1 - 1$. Then $\sum_{i=1}^{h} |X_i| \leqslant r(k_1 - 1)$ which implies that $\sum_{j=1}^{p} |Y_j| \geqslant r$ or $p \geqslant r/(t-1)$. Since vertices in distinct groups (one vertex per group) in G and the corresponding vertices in G_1 induce isomorphic subgraphs, the vertices $v_1, \ldots, v_p$ induce a subgraph of G_1 with property Π (since Π is hereditary). Since this subgraph has property Π, it does not contain H_s as an induced subgraph. We choose r large enough so that any graph on $r/(t-1)$ vertices that does not contain a clique of size $|H_s|$ has an independent set of size k_1. With this choice of r, it follows that G_1 does contain an independent set of size k_1. The number r depends only on $|H_s|$ and the parameter k_1 and not on $n_1 = |V(G_1)|$. So the reduction is achieved in $O(f(k_1){n_1}^{\alpha})$ time where f is some function of k_1 and α is some fixed constant independent of k_1. □

Next, we will show that the problem is $W[1]$-hard even if none of the graphs in the finite forbidden set is complete-bipartite.

Theorem 5. *Let Π be a hereditary property that includes all trivial graphs but not all complete graphs, and that has a finite forbidden set $\mathscr{F} = \{H_1, H_2, \ldots, H_s\}$. Then the problem $P(G, k, \Pi)$ is $W[1]$-complete.*

Proof. The fact that the problem is in $W[1]$ has already been proved in Lemma 3. Assume that none of the graphs H_i in the forbidden set of Π is complete-bipartite. Let H_s be the clique of smallest size that is not in Π, hence in the forbidden set $\mathscr{F}$.

For a graph H_i in $\mathscr{F}$, select (if possible) a subset of vertices Z such that the vertices in Z are independent and every vertex in Z is connected to every vertex in $H_i \backslash Z$. Let $\{H_{ij} \mid 1 \leqslant j \leqslant s_i\}$ be the set of graphs obtained from H_i by removing such a set Z for every possible choice of Z. Since H_i is not complete-bipartite, every H_{ij} is a nontrivial graph. Let $\mathscr{F}_1 = \mathscr{F} \cup \{H_{ij} \mid 1 \leqslant i \leqslant s, 1 \leqslant j \leqslant s_i\}$. Note that $\mathscr{F}_1$ contains a clique of size $|H_s| - 1$ because a set Z, consisting of a single vertex, can be deleted from the clique H_s. Let Π_1 be the hereditary property defined by the forbidden set $\mathscr{F}_1$. Observe that Π_1 also includes all trivial graphs but not all complete graphs. Let P_1 be the problem $P(G_1, k_1, \Pi_1)$.

We will prove that P_1 is $W[1]$-hard later. Now, we will reduce P_1 to the problem $P(G,k,\Pi)$ at hand.

Given G_1, we construct a graph G as follows. Let $V(G)=V(G_1)\cup D$ where D is a set of r independent vertices (r to be specified later). Every vertex in $V(G_1)$ is connected to every vertex in D. Let $v=\max_i(|H_i|)$.

We claim that G_1 has an induced subgraph of size k_1 with property Π_1 if and only if G has k_1+r vertices that induce a subgraph with property Π.

Let A be a subset of $V(G_1)$, $|A|=k_1$ such that $G_1[A]\in\Pi_1$. Let $S=A\cup D$. Suppose on the contrary that $G[S]$ contains some H_i as a subgraph. If this H_i contains some vertices from D, we throw away these independent vertices. The remaining portion of H_i, which is some H_{ij}, $1\leqslant j\leqslant s_i$, must lie in $G[A]$. But this is a contradiction because $G[A]=G_1[A]$ and by hypothesis, $G_1[A]$ has property Π_1 and it cannot contain any H_{ij}. Similarly, H_i cannot lie entirely in $G[A]$ because $\mathscr{F}\subseteq\mathscr{F}_1$, so $G[A]$ does not contain any H_i as induced subgraph. Therefore $G[S]$ does not contain any H_i, hence it has property Π and $|S|=k_1+r$.

Conversely, suppose we can choose a set S, $|S|=k_1+r$ such that $G[S]$ does not contain any H_i. Since $|D|=r$ we must choose at least k_1 vertices from $V(G_1)$. Let $A\subseteq S\cap V(G_1)$ with cardinality k_1. If $G[A]$ does not contain any H_{ij}, we are through. Otherwise let $H_{i_0j_0}\subseteq G[A]$ for some i_0,j_0. Now $H_{i_0j_0}$ is obtained from H_{i_0} by deleting an independent set of size at most v. Hence S can contain at most $v-1$ vertices from D, otherwise we could add sufficient number of vertices from D to the graph $H_{i_0j_0}$ to get a copy of H_{i_0} which is not possible. Hence, $|S\cap D|<v$ which implies that $|S\cap V(G_1)|>k_1+r-v$. Thus, $G_1[S\cap V(G_1)]$ is an induced subgraph of G_1 of size at least k_1+r-v that does not contain any H_i, in particular it does not contain H_s which is a clique of size say μ. We can select r (as before, by Ramsey theorem) such that any graph on k_1+r-v vertices that does not contain a μ-clique has an independent set of size k_1. Hence, G_1 has an independent set of size k_1 which has property Π_1. The number r depends only on the family $\mathscr{F}$ and parameter k_1 and not on $n_1=|V(G_1)|$. So the reduction is achieved in $O(g(k_1){n_1}^{\beta})$ time where g is some function of k_1 and β is a constant. Also $|V(G)|=|V(G_1)|+r$, so the size of the input problem increases only by a constant.

We argue that the problem P_1 is $W[1]$-hard. If any of the H_{ij} is complete-bipartite, then this follows from Lemma 4. Otherwise, we repeatedly apply the construction given at the beginning of the proof, of removing a set Z of vertices from each graph in the forbidden set, to get families $\mathscr{F}_2,\mathscr{F}_3,\ldots$ and corresponding problems $P_2,P_3,\ldots$ such that there is a parametric reduction from P_{m+1} to P_m. Since $\mathscr{F}_{m+1}$ contains a smaller clique than the smallest clique in $\mathscr{F}_m$, eventually some family $\mathscr{F}_{m_0}$ contains a clique of size 2 (the graph K_2) or a complete-bipartite graph. In the former case, the problem P_{m_0} is the same as the parameterized independent set problem, and so is $W[1]$-hard. In the latter case P_{m_0} is $W[1]$-hard by Lemma 4. $\square$

4.2. Properties with infinite forbidden sets

Here we extend Theorem 5 to the case when the forbidden set is infinite.

Theorem 6. *Let Π be a hereditary property that includes all trivial graphs but not all complete graphs* (*or vice versa*). *Then the problem $P(G,k,\Pi)$ is $W[1]$-complete.*

Proof. Every hereditary property is defined by a (possibly infinite) forbidden set [11]; let the forbidden family for Π be $\mathscr{F}$. The proof is almost the same as in Theorem 5. Note that Lemma 4 does not depend on finiteness of the forbidden family. Also the only point where the finiteness of $\mathscr{F}$ is used in Theorem 5 is in the argument that if $G[A]$ does contain some $H_{i_0j_0}$ then S can contain at most $v-1$ vertices from the set D. This argument can be modified as follows. Since $G[A]$ contains some $H_{i_0j_0}$, $|V(H_{i_0j_0})|\leqslant|A|=k_1$. Also $H_{i_0j_0}$ is obtained from some H_i by removing an independent set adjacent to all other vertices of H_i. (If there are more than one such H_i from which $H_{i_0j_0}$ is obtained, we choose an arbitrary H_i.) Let $v_1=\max(|V(H_i)|-|V(H_{ij})|)$ where the maximum is taken over all H_{ij} such that $|V(H_{ij})|\leqslant k_1$. Hence, if $G[A]$ does contain some $H_{i_0j_0}$, we can add at most v_1 vertices from D to get H_{i_0}. So S must contain less than v_1 vertices from D. The choice of r will have to be modified accordingly. □

Corollary 7 follows from Theorem 6 since the collection of forests is a hereditary property with the forbidden set as the set of all cycles. This collection includes all trivial graphs and does not include any complete graph on more than two vertices.

Corollary 7. *The following problem is $W[1]$-complete*:
Given (G,k), does G have k vertices that induce a forest?

This problem is the parametric dual of the Undirected Feedback Vertex Set problem which is known to be fixed parameter tractable [5].

Corollary 8. *The following problem is $W[1]$-complete*:
Given (G,k), does there exist an induced subgraph of G with k vertices that is bipartite?

Proof. Completeness follows from Theorem 6 since all independent sets are bipartite and no clique of size at least 3 is bipartite.

We give here a more direct proof that the problem is in $W[1]$ by reducing the problem to the weight k satisfying the assignment problem. This problem, which asks whether a given boolean formula has a satisfying assignment whose weight (the number of ones) is k, has been shown to be in $W[1]$ [5].

Given the graph G, consider the boolean formula:

$$\bigwedge_{u\in V(G)}(\overline{x_u}\vee\overline{y_u})\bigwedge_{(u,v)\in E(G)}((\overline{x_u}\vee\overline{x_v})\wedge(\overline{y_u}\vee\overline{y_v})).$$

We claim that G has an induced bipartite subgraph of size k if and only if the above formula has a satisfying assignment with weight k. Suppose G has an induced bipartite subgraph with k vertices with partition V_1 and V_2. Now for each vertex in V_1 assign $x_u=1,\ y_u=0$, for each vertex in V_2 assign $x_u=0,\ y_u=1$ and assign $x_u=y_u=0$ for the remaining vertices. It is easy to see that this assignment is a weight k satisfying assignment for the above formula.

Conversely, if the above formula has a weight k satisfying assignment, the vertices u such that $x_u = 1,\ y_u = 0$ or $x_u = 0,\ y_u = 1$ induce a bipartite subgraph of G with k vertices. □

Corollary 9 can be proved along similar lines of Corollary 8.

Corollary 9. *The following problem is* $W[1]$*-complete*:
Given (G,k) *and a constant* l, *does there exist an* l*-colorable induced subgraph of size* k?

Finally, we address the parametric dual of the problem addressed in Corollary 8. Given a graph G, and an integer k, are there k vertices in G whose removal makes the graph bipartite? We will call this problem '$n - k$ bipartite'.

The precise parameterized complexity of this problem is unknown. Although bipartiteness is a hereditary property, it has an infinite forbidden set, and so the problem is not covered by Cai's result [1].

The 'edge' counterpart of the problem, given a graph G with m edges, and an integer k, are there k edges whose removal makes the graph bipartite, is the same as asking for a cut in the graph of size $m - k$. It is known [12] that there exists a parameterized reduction from this problem to the following problem, which we call 'all but k 2-SAT'.

Given: A boolean 2 CNF formula F.

Parameter: An integer k.

Question. Is there an assignment to the variables of F that satisfies all but at most k clauses of F?

We show the following.

Theorem 10. *There is a parameterized reduction from the '*$n - k$ *bipartite problem' to the 'all but* k *2-SAT' problem.*

Proof. Given a graph G, for every vertex, we set two variables (x_u, y_u) and construct clauses in the same manner as in the proof of Corollary 8. The clauses are as follows:

Set 1:

$$\overline{x_u} \vee \overline{y_u} \quad \forall u \in V(G),$$

$$\overline{x_u} \vee \overline{x_v}; \overline{y_u} \vee \overline{y_v} \quad \forall (u,v) \in E(G).$$

Each clause in Set 1 is repeated $k + 1$ times.

Set 2: $x_u \vee y_u\ \forall u \in V(G)$.

We claim that it is possible to delete k vertices to make the given graph bipartite if and only if there is an assignment to the variables in the above formula that makes all but at most k clauses true.

If there is an assignment that makes all but at most k clauses true, then the clauses in Set 1 must be true because each of them occurs $k+1$ times. This ensures that the variables x_u, y_u corresponding to the vertices are assigned, respectively, $0,0$ or $0,1$ or $1,0$ and each edge $e=(u,v)$ has $x_u=x_v=y_u=y_v=0$ or $x_u=0,\ y_u=1$ and $x_v=1$, $y_u=0$ or vice versa. The vertices s for which $x_s=y_s=0$ are deleted to get a bipartite graph. At most k clauses in Set 2 are false. This ensures that at most k vertices are deleted.

Conversely, if there exist k vertices whose removal results in a bipartite graph with partition $V_1 \cup V_2$, consider the assignment corresponding to each vertex u in the graph, $x_u=y_u=0$ if the vertex u is deleted, $x_u=1,\ y_u=0$ if $u \in V_1$ and $x_u=0,\ y_u=1$ if $u \in V_2$.

It is easy to see that this assignment makes all but at most k clauses of the formula true.

Note that the reduction is actually a polynomial time reduction. □

5. Concluding remarks

We have characterized the hereditary properties for which finding an induced subgraph with k vertices having the property in a given graph is $W[1]$-complete. In particular, using Ramsey theorem, we have shown that if the property includes all trivial graphs and all complete graphs or if it excludes some trivial graph as well as some complete graph, then the problem is fixed parameter tractable and is $W[1]$-complete otherwise. For some of these specific properties, more efficient fixed parameter algorithms (not based on Ramsey numbers) should be possible.

It remains an open problem to determine the parameterized complexity of both of the problems stated in Theorem 10 (the '$n-k$ bipartite problem' and the 'all but k 2-SAT' problem). More generally, the parameterized complexity of the node-deletion problem for a hereditary property with an infinite forbidden set is open.

We remark that our results prove that the parametric dual of a problem considered by Cai [1] (and proved FPT) is $W[1]$-complete. This observation adds weight to the observation (first made in [12]) that typically parametric dual problems have complimentary parameterized complexity. It would be interesting to identify some general conditions that guarantee such complimentary parameterized complexity for parametric dual problems.

Finally, the NP-hard results of the problems addressed in the paper have edge counterparts (are there k edges that induce a subgraph with property Π?) and generalizations to directed graphs [14,11]. It would be interesting to explore the parameterized complexity of those generalizations.

Acknowledgements

We thank Mike Fellows for his comments on an earlier draft of the paper.

References

[1] L. Cai, Fixed parameter tractability of graph modification problem for hereditary properties, Inform. Process. Lett. 58 (1996) 171–176.

[2] M. Cesati, Perfect code is $W[1]$-complete, Inform. Process. Lett. 81 (2002) 163–168.

[3] R.G. Downey, M.R. Fellows, Fixed parameter tractability and completeness I: basic theory, SIAM J. Comput. 24 (1995) 873–921.

[4] R.G. Downey, M.R. Fellows, Fixed parameter tractability and completeness II: completeness for $W[1]$, Theoret. Comput. Sci. 141 (1995) 109–131.

[5] R.G. Downey, M.R. Fellows, Parameterized Complexity, Springer, New York, 1999.

[6] R.G. Downey, M.R. Fellows, V. Raman, The complexity of irredundant sets parameterized by size, Discrete Appl. Math. 100 (2000) 155–167.

[7] M.R. Fellows, C. McCartin, On the parametric complexity of schedules to minimize tardy tasks, Theoret. Comput. Sci., to appear.

[8] M.R. Garey, D.S. Johnson, Problem [GT21], Computers and Intractability, A Guide to The Theory of NP-Completeness, Freeman and Company, New York, 1979.

[9] M.C. Golumbic, Algorithmic Graph Theory and Perfect Graphs, Academic Press, New York, 1980.

[10] Harary, Graph Theory, Addison-Wesley Publishing Company, Reading, MA, 1969.

[11] J.M. Lewis, M. Yannakakis, The Node-Deletion Problem for Hereditary Properties is NP-complete, J. Comput. System Sci. 20 (2) (1980) 219–230.

[12] M. Mahajan, V. Raman, Parameterizing above Guaranteed Values: MaxSat and MaxCut, J. Algorithms 31 (1999) 335–354.

[13] A. Natanzon, R. Shamir, R. Sharan, Complexity Classification of Some Edge Modification Problems, in Proc. 25th Workshop on Graph-Theoretic Concepts in Computer Science (WG'99), Ascona, Switzerland, Lecture Notes in Computer Science, Springer, Berlin, 1999.

[14] M. Yannakakis, Edge-deletion problems, SIAM J. Comput. 10 (2) (1981) 297–309.